



Release 3.1a John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MPsrch\_pp protein - protein database search, using Smith-Waterman algorithm  
Run on: Sat May 13 10:22:35 2000; Maspar time 3.95 Seconds  
Tabular output not generated. 258.173 Million cell updates/sec

Title: >US-09-331-631-5  
Description: (33-75). From US09331631.pep (2 of 4)  
Perfect Score: 343  
Sequence: 1 NOEDPOTECOCORRCROESDPDPOOYCORCKEICEEERY 43

Scoring table: PAM 150  
Gap 11

Searched: 188963 seqs, 23686106 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: a:geneseq35  
1:geneseqp

Statistics: Mean 22.157; Variance 99.354; scale 0.223

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	343	100.0	625	1	Macadamia integrifolia	1.45e-22
2	337	98.3	625	1	Macadamia integrifolia	5.15e-22
3	335	97.7	666	1	Macadamia integrifolia	7.86e-22
4	155	45.2	525	1	Theobroma cacao antimi	1.09e-05
5	155	45.2	566	1	Sequence encoded by 67	1.09e-05
6	135	39.4	590	1	Gossypium hirsutum ant	5.57e-04
7	119	34.7	28	1	Stenocarpus sinuatus a	1.22e-02
8	115	33.5	593	1	Zea mays antimicrobial	8.17e-02
9	109	31.8	33	1	Zea mays antimicrobial	4.40e-01
10	100	29.2	35	1	Antimicrobial maize pe	3.96e+00
11	88	25.7	625	1	Peanut allergen Ara hi	6.80e+00
12	85	24.6	637	1	Hordeum vulgare antimi	9.73e+00
13	83	24.2	106	1	AcanaP24.	9.73e+00
14	83	24.2	107	1	AcanaP23.	9.73e+00
15	83	24.2	441	1	Steroid hormone recept	9.73e+00
16	83	24.2	441	1	Peroxisome proliferato	9.73e+00
17	83	24.2	614	1	Arachis hypogaea antim	9.73e+00
18	83	24.2	614	1	Peanut allergen Ara hi	9.73e+00
19	79	23.0	439	1	R33745	1.98e+01
20	79	23.0	440	1	Mouse peroxisome proli	1.98e+01
21	79	23.0	440	1	Peroxisome proliferato	1.98e+01
22	79	23.0	919	1	Human androgen recepto	1.98e+01
23	79	23.0	919	1	Androgen receptor.	1.98e+01

24	76	22.2	516	1	P61362	Soybean glycinin A3B4	3.35e+01
25	76	22.2	918	1	R12223	Human androgen recepto	3.98e+01
26	75	21.9	204	1	W59609	DNA-binding/dimerizati	3.98e+01
27	75	21.9	468	1	R74053	Human peroxisome proli	3.98e+01
28	75	21.9	768	1	R27689	Rabbit beta-8 intergri	3.98e+01
29	74	21.6	55	1	W80489	Murine vascular endothe	4.74e+01
30	74	21.6	55	1	W04825	Vascular endothelial g	4.74e+01
31	74	21.6	145	1	W86213	Human VEGF-B truncated	4.74e+01
32	74	21.6	147	1	W86212	Human VEGF-B truncated	4.74e+01
33	74	21.6	150	1	W86211	Human VEGF-B truncated	4.74e+01
34	74	21.6	152	1	W86210	Human VEGF-B truncated	4.74e+01
35	74	21.6	155	1	W86209	Human VEGF-B truncated	4.74e+01
36	74	21.6	160	1	W86208	Human VEGF-B truncated	4.74e+01
37	74	21.6	167	1	W86234	Human VEGF-B full leng	4.74e+01
38	74	21.6	188	1	W86201	Human vascular endothe	4.74e+01
39	74	21.6	188	1	W04826	Heart vascular endothe	4.74e+01
40	74	21.6	188	1	W80490	Murine vascular endothe	4.74e+01
41	74	21.6	190	1	R91710	AcenaP4.	4.74e+01
42	74	21.6	195	1	W80491	Murine vascular endothe	4.74e+01
43	74	21.6	195	1	W04827	Heart vascular endothe	4.74e+01
44	74	21.6	176	1	W50893	Human laminin BI chain	4.74e+01
45	73	21.3	188	1	W08064	Murine VWF167.	5.63e+01

## ALIGNMENTS

RESULT	1	W62830 standard; Protein; 625 AA.
ID	W62830	
AC	27-OCR-1998 (first entry)	
DE	Macadamia integrifolia antimicrobial protein.	
OS	antimicrobial protein; infestation; control.	
KW	Macadamia integrifolia.	
FT	Key	Location/Qualifiers
FT	Peptide	1..28
FT	Protein	/note="signal peptide"
FT		29..666
FT		/note="mature protein"
PN	W09827805-A1.	
PD	02-JUL-1998.	
PE	22-DEC-1997; AU00874.	
PR	20-DEC-1996; AU004275.	
PA	(RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.	
PI	Bower NI, Goulter KC, Grefn JL, Manners JM, Marcus JP;	
DR	WPI; 98-377279/32.	
DT	N-PSDB; V42316.	
PT	Novel anti-microbial protein from e.g. Macadamia integrifolia -	
PT	useful for controlling microbial infestations of plants or mammals	
PS	Claim 1: Page 43-45; 96pp English.	
CC	The sequence is that of an antimicrobial protein which can	
CC	be used to control microbial infestations in plants and mammalian	
CC	animals.	
CC	Sequence 625 AA:	
SO		
Query Match	100.0%; Score 343; DB 1; Length 625;	
Best local Similarity	100.0%; Pred. No. 1.45e-22;	
Matches	43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
DB	33 NOEDPOTECOCORRCROESDPDPOOYCORCKEICEEERY 75	
QY	33 NOEDPOTECOCORRCROESDPDPOOYCORCKEICEEERY 75	
RESULT	2	
ID	W62828 standard; Protein; 666 AA.	
AC	W62828;	
DE	27-OCR-1998 (first entry)	
DE	Macadamia integrifolia antimicrobial protein.	
KW	antimicrobial protein; infestation; control.	
OS	Macadamia integrifolia.	
FT	Key	Location/Qualifiers
FT	Peptide	1..28
FT		/note="signal peptide"

FT Protein 29..666  
 /note="mature protein"  
 PN MO9827805-A1.  
 PD 02-JUL-1998.  
 PE 22-DEC-1997; AU0874.  
 PR 20-DEC-1996; AU-004275.  
 PA (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.  
 PI Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP;  
 DR WPI: 98-377279/32.  
 N-PSDB: V42310.  
 PT Novel anti-microbial protein from e.g. Macadamia integrifolia -  
 useful for controlling microbial infestations of plants or mammals  
 PS Claim 1; Page 34-36; 96pp; English.  
 CC The sequence is that of an antimicrobial protein which can  
 be used to control microbial infestations in plants and mammalian  
 CC animals.  
 CC  
 SO Sequence 666 AA;

Query Match 98.3%; Score 337; DB 1; Length 666;  
 Best Local Similarity 97.7%; Pred. No. 5.15e-22;  
 Matches 42; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

DB 74 NOEDPOTECQCCQRCRQOESGPRQOQYCCQRCRKEICEEERY 116  
 QY 33 NOEDPOTECQCCQRCRQOESDPQOQYCCQRCRKEICEEERY 75

RESULT 3  
 ID W62829 standard; Protein; 666 AA.  
 AC W62829.  
 DT 27-OCT-1998 (first entry)  
 DE Macadamia integrifolia antimicrobial protein.  
 KW antimicrobial protein; infestation; control.  
 OS Macadamia integrifolia.  
 FH Key Location/Qualifiers  
 FT Peptide 1..28  
 /note="signal peptide"  
 FT Protein 29..666  
 /note="mature protein"  
 PN MO9827805-A1.  
 PD 02-JUL-1998.  
 PE 22-DEC-1997; AU0874.  
 PR 20-DEC-1996; AU-004275.  
 PA (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.  
 PI Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP;  
 DR WPI: 98-377279/32.  
 N-PSDB: V42311.  
 PT Novel anti-microbial protein from e.g. Macadamia integrifolia -  
 useful for controlling microbial infestations of plants or mammals  
 PS Claim 1; Page 39-41; 96pp; English.  
 CC The sequence is that of an antimicrobial protein which can  
 be used to control microbial infestations in plants and mammalian  
 CC animals.  
 CC  
 SO Sequence 666 AA;

Query Match 97.7%; Score 335; DB 1; Length 666;  
 Best Local Similarity 93.0%; Pred. No. 7.86e-22;  
 Matches 40; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

DB 74 NOEDPOTECQCCQRCRQOESGPRQOQYCCQRCRKEICEEERY 116  
 QY 33 NOEDPOTECQCCQRCRQOESDPQOQYCCQRCRKEICEEERY 75

RESULT 4  
 ID W62831 standard; Protein; 525 AA.  
 AC W62831.  
 DT 27-OCT-1998 (first entry)  
 DE Theobroma cacao antimicrobial protein.  
 KW antimicrobial protein; infestation; control.  
 OS Theobroma cacao.  
 PN MO9827805-A1.  
 PD 02-JUL-1998.

PF 22-DEC-1997; AU0874.  
 PR 20-DEC-1996; AU-004275.  
 PA (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.  
 PI Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP;  
 DR WPI: 98-377279/32.  
 PT Novel anti-microbial protein from e.g. Macadamia integrifolia -  
 useful for controlling microbial infestations of plants or mammals  
 PS Claim 1; Page 47-49; 96pp; English.  
 CC The sequence is that of an antimicrobial protein which can  
 be used to control microbial infestations in plants and mammalian  
 CC animals.  
 CC  
 SO Sequence 525 AA;

Query Match 45.2%; Score 155; DB 1; Length 525;  
 Best Local Similarity 47.5%; Pred. No. 1.09e-05;  
 Matches 19; Conservative 9; Mismatches 12; Indels 0; Gaps 0;

DB 78 EEEIQRQYQCCQRCRQOESGPRQOQYCCQRCRKEICEEERY 117  
 QY 34 QEDPOTECQCCQRCRQOESDPQOQYCCQRCRKEICEEERY 73

RESULT 5  
 ID R20181 standard; Protein; 566 AA.  
 AC R20181.  
 DT 16-APR-1992 (first entry)  
 DE Sequence encoded by 67 kD T. cacao protein cDNA.  
 KW Cocoa; flavour; vicillin; seed storage protein.  
 OS Theobroma cacao.  
 PN WO9119801-A.  
 PD 26-DEC-1991.  
 PE 07-JUN-1991; G00914.  
 PR 11-JUN-1990; GB-013016.  
 PA (MRSC ) MARS UK LTD.  
 PI Spencer ME, Hodge R, Deakin EA, Ashton S;  
 DR WPI: 92-024418/03.  
 N-PSDB: Q20377.  
 PT Recombinant cocoa proteins - are responsible for flavour in cocoa  
 PT beans and produced in large quantities using yeast and bacterial  
 PT expression vectors  
 PS Claim 4; Fig 2; 59pp; English.  
 CC The inventors claim a 67 kD and 31 kD T. cacao protein, and  
 CC fragments, and encoding DNAs. The 47 kD and 31 kD proteins are  
 CC derived from the 67 kD precursor. T. cacao protein cDNA was  
 CC detected in a cDNA library prepared from immature cocoa beans RNA  
 CC using a probe based on the AA sequence of a CNBR peptide common to  
 CC the 47 kD and 31 kD polypeptides. Homology searches revealed close  
 CC homologues between the 67 kD polypeptide and the vicillins, which are  
 CC seed storage proteins.  
 CC  
 SO Sequence 566 AA;

Query Match 45.2%; Score 155; DB 1; Length 566;  
 Best Local Similarity 47.5%; Pred. No. 1.09e-05;  
 Matches 19; Conservative 9; Mismatches 12; Indels 0; Gaps 0;

DB 78 EEEIQRQYQCCQRCRQOESGPRQOQYCCQRCRKEICEEERY 117  
 QY 34 QEDPOTECQCCQRCRQOESDPQOQYCCQRCRKEICEEERY 73

RESULT 6  
 ID W62832 standard; Protein; 590 AA.  
 AC W62832.  
 DT 27-OCT-1998 (first entry)  
 DE Gossypium hirsutum antimicrobial protein.  
 KW antimicrobial protein; infestation; control.  
 OS Gossypium hirsutum.  
 PN WO9827805-A1.  
 PD 02-JUL-1998.  
 PE 22-DEC-1997; AU0874.  
 PR 20-DEC-1996; AU-004275.  
 PA (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.  
 PI Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP;







